

VU Research Portal

Deletion of two exons from the *Lymnaea stagnalis* β 164-N-acetylglucosaminyltransferase gene elevated the kinetic efficiency of the encoded enzyme for both UDP-sugar donor and acceptor substrates

Bakker, H.; van Tetering, A.; Agterberg, L.; Smit, A.B.; van den Eijnden, D.H.; van Die, I.M.

published in

Journal of Biological Chemistry

1997

DOI (link to publisher)

[10.1074/jbc.272.30.18580](https://doi.org/10.1074/jbc.272.30.18580)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Bakker, H., van Tetering, A., Agterberg, L., Smit, A. B., van den Eijnden, D. H., & van Die, I. M. (1997). Deletion of two exons from the *Lymnaea stagnalis* β 164-N-acetylglucosaminyltransferase gene elevated the kinetic efficiency of the encoded enzyme for both UDP-sugar donor and acceptor substrates. *Journal of Biological Chemistry*, 272, 18580-18585. <https://doi.org/10.1074/jbc.272.30.18580>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Deletion of Two Exons from the *Lymnaea stagnalis* β 1 \rightarrow 4-*N*-Acetylglucosaminyltransferase Gene Elevates the Kinetic Efficiency of the Encoded Enzyme for Both UDP-sugar Donor and Acceptor Substrates*

(Received for publication, November 13, 1996, and in revised form, March 27, 1997)

Hans Bakker[‡], Angelique Van Tetering[‡], Marja Agterberg[‡], August B. Smit[§],
Dirk H. Van den Eijnden[‡], and Irma Van Die[‡]¶

From the Departments of [‡]Medical Chemistry and [§]Experimental Zoology, Vrije Universiteit,
1081 BT Amsterdam, The Netherlands

Lymnaea stagnalis UDP-GlcNAc:GlcNAc β -R β 1 \rightarrow 4-*N*-acetylglucosaminyltransferase (β 4-GlcNAcT) is an enzyme with structural similarity to mammalian UDP-Gal:GlcNAc β -R β 1 \rightarrow 4-galactosyltransferase (β 4-GalT). Here, we report that also the exon organization of the genes encoding these enzymes is very similar. The β 4-GlcNAcT gene (12.5 kilobase pairs, spanning 10 exons) contains four exons, encompassing sequences that are absent in the β 4-GalT gene. Two of these exons (exons 7 and 8) show a high sequence similarity to part of the preceding exon (exon 6), suggesting that they have originated by exon duplication. The exon in the β 4-GalT gene, corresponding to β 4-GlcNAcT exon 6, encodes a region that has been proposed to be involved in the binding of UDP-Gal. The question therefore arose, whether the repeating sequences encoded by exon 7 and 8 of the β 4-GlcNAcT gene would determine the specificity of the enzyme for UDP-GlcNAc, or for the less preferred UDP-GalNAc. It was found that deletion of only the sequence encoded by exon 8 resulted in a completely inactive enzyme. By contrast, deletion of the amino acid residues encoded by exons 7 and 8 resulted in an enzyme with an elevated kinetic efficiency for both UDP-sugar donors, as well as for its acceptor substrates. These results suggest that at least part of the donor and acceptor binding domains of the β 4-GlcNAcT are structurally linked and that the region encompassing the insertion contributes to acceptor recognition as well as to UDP-sugar binding and specificity.

Glycosyltransferases form a large family of functionally related, membrane-bound enzymes that are involved in the biosynthesis of the carbohydrate moieties of glycoproteins and glycolipids (1, 2). Recently we have identified a novel glycosyltransferase by the isolation of a UDP-GlcNAc:GlcNAc β -R β 1 \rightarrow 4-*N*-acetylglucosaminyltransferase (β 4-GlcNAcT)¹ cDNA from the prostate gland of the snail *Lymnaea stagnalis* (3). In

vitro, the recombinant β 4-GlcNAcT catalyzes the transfer of GlcNAc from UDP-GlcNAc in β 1 \rightarrow 4 linkage to various β -*N*-acetylglucosaminides (3, 4). The β 4-GlcNAcT cDNA appeared to show a significant sequence similarity to the mammalian UDP-Gal:GlcNAc β -R β 1 \rightarrow 4-galactosyltransferase (β 4-GalT) cDNAs, with an overall resemblance between the predicted amino acid sequences of about 30% (3, 5–7). Based on the genetic and enzymatic relationship of the β 4-GlcNAcT and the β 4-GalTs, we have suggested that these enzymes constitute a separate glycosyltransferase gene family, the members of which are capable of catalyzing the transfer of a specific sugar from their respective UDP-sugar donors in a β 1 \rightarrow 4 linkage toward a terminal β -linked GlcNAc residue in the acceptor (3, 8). Based on enzymatic properties, we have proposed that also UDP-GalNAc:GlcNAc β -R β 1 \rightarrow 4-*N*-acetylgalactosaminyltransferase (β 4-GalNAcT), detected in several non-vertebrate species (9–13), belongs to this family (14). The primary structure of this enzyme, however, is still unknown.

The reaction catalyzed by glycosyltransferases typically involves two substrates and often a divalent cation cofactor. This suggests that the enzymes consist of several functional domains involved in substrate and cofactor binding, respectively. Comparison of the conserved and variable regions of genetically related glycosyltransferases with different properties would open possibilities to address structure-function relationships. As snails and mammals are evolutionary distant species, comparison of the genomic organization of the genes that encode these enzymes might give insight in their way of divergence, that resulted in genes encoding enzymes with a different UDP-sugar specificity. The genomic organization of the murine and human β 4-GalT genes have been described previously (15, 16). Here we report the organization of the *L. stagnalis* gene that codes for the β 4-GlcNAcT. The intron-exon distribution was determined and compared with that of the β 4-GalT gene. Mutant β 4-GlcNAcT cDNAs were constructed by deletion of sequences that do not have a counterpart in the β 4-GalT gene, and expressed in COS cells. Comparison of the kinetic parameters of the resulting mutant and parental enzymes showed that the insertion in the β 4-GlcNAcT and its surrounding regions contribute to acceptor recognition as well as to UDP-sugar binding and specificity.

EXPERIMENTAL PROCEDURES

Materials, Bacterial Strains, and Cells—Acceptor substrates were obtained from Sigma (compounds 1, 2, and 3) and from Toronto Research Chemicals (compounds 5 and 6). Compound 4 was a gift of Dr. O.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom reprint requests should be addressed: Dept. of Medical Chemistry, Vrije Universiteit, Van der Boerhorststraat 7, 1081 BT Amsterdam, The Netherlands. Tel.: 31-20-4448157; Fax: 31-20-4448144; E-mail: im.van_die.medchem@med.vu.nl.

¹ The abbreviations used are: β 4-GlcNAcT, UDP-GlcNAc:GlcNAc β -R β 1 \rightarrow 4-*N*-acetylglucosaminyltransferase; β 4-GalNAcT, UDP-GalNAc:GlcNAc β -R β 1 \rightarrow 4-*N*-acetylgalactosaminyltransferase; β 4-GalT, UDP-Gal:GlcNAc β -R β 1 \rightarrow 4-galactosyltransferase; kb, kilobase pair(s); PCR, polymerase chain reaction; HPAEC, high pH anion-exchange chroma-

tography; PAD, pulsed amperometric detection; pNP, *para*-nitrophenyl; WFA, *W. floribunda* lectin.

Hindsgaul, University of Alberta, Alberta, Canada. UDP-[^3H]GlcNAc, UDP-[^3H]GalNAc and UDP-[^3H]Gal were purchased from NEN Life Science Products, and were diluted with unlabeled UDP-sugars (Sigma) to the desired specific radioactivity.

Recombinant plasmids were propagated in the *Escherichia coli* K12 strain XL1-Blue (Stratagene). COS-7M6 cells (ATCC 1651) were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, and 1:50 penicillin/streptomycin solution (all from Life Technologies, Inc.). Synthetic oligonucleotides were obtained from Isogen Bioscience BV (Maarssen, the Netherlands).

DNA Techniques—The genomic clone $\lambda 5$ has been isolated previously (3) from a genomic λ EMBL3A library of *L. stagnalis* (17). Isolation of plasmid DNA was carried out by a modification of the miniprep method, as described in Ref. 18. Plasmids used for transfection of COS cells were isolated by the Qiagen plasmid protocol, using a QIAGEN-tip 100 minicolumn. Restriction enzymes and other DNA-modifying enzymes were used according to the manufacturer. Dideoxynucleotide chain-terminating sequencing reactions (19) were performed on double-stranded plasmid DNA, with the T7 DNA sequencing kit (Pharmacia Biotech Inc.), [α - ^{32}P]dATP (Amersham), using M13 universal primer, the KS and SK primers, and several sequence-specific synthetic oligonucleotide primers. Southern blotting was performed as described previously (3). PCR with sequence-specific primers was performed using Ulta-polymerase (Perkin-Elmer) by 25 cycles (1 min at 95 °C, 1 min at 63 °C, 1 min at 72 °C). For cloning purposes, amplified fragments were subsequently purified according to the QIAquick PCR purification protocol (Qiagen Inc.).

Construction of pPROTA Hybrid Plasmids—The plasmid pMC135, containing a fusion between part of the protein A sequence and $\beta 4$ -GlcNAcT cDNA, was constructed as follows. A *Bam*HI-*Eco*RI adapter was ligated in the *Bam*HI site of pVTBac-P11.4, carrying 5'-truncated $\beta 4$ -GlcNAcT cDNA (3). The resulting 1.4-kb *Eco*RI fragment from this construct was ligated into an *Eco*RI-digested pPROTA vector (20). Plasmids carrying mutant $\beta 4$ -GlcNAcT genes were derived from pMC135 by exchange of the 0.52-kb *Xho*I-*Bgl*II fragment for a PCR fragment carrying the desired deletion. For PCR of the deletion fragments, a sense primer (bases 439–456 of the $\beta 4$ -GlcNAcT cDNA; Ref. 3) and the antisense primer ID8 (ATAGATCTTAAATCTGTCCGGGTTTCACAT-TCCAG) or ID16 (ATAGATCTTAAATCTGTTCGGGTCACGTTTC) was used. The antisense primer ID8, used for construction of pMC142, consists of a part complementary to the 3' end of exon 6, and a part (11 base pairs), complementary to the 5' end of exon 9 (*Bgl*II restriction site underlined). The antisense primer ID16, used for construction of pMC166, consists of a part complementary to the 3' end of exon 7, and the same 5' part of exon 9 as ID8. PCR fragments obtained with these primers were digested with *Xho*I and *Bgl*II, and ligated into *Xho*I-*Bgl*II-digested pMC135. After transformation and plasmid isolation of several transformants, the desired plasmids (pMC142 encoding protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8 and pMC166 encoding protA- $\beta 4$ -GlcNAcT $\Delta 8$) were selected by size determination of the internal *Sal*I fragment followed by determination of the nucleotide sequence of the complete *Xho*I-*Bgl*II fragment with sequence-specific primers.

Expression in COS-7M6 Cells—Recombinant pPROTA chimeric constructs were transiently transfected to COS cells (3×10^5 cells/10-cm dish), using the calcium phosphate precipitation technique as described previously (21); after 24 h, the medium was replaced by fresh medium; medium was harvested at 48, 72, and 96 h after transfection and each time replaced by fresh medium. The harvested media were pooled, stored at –20 °C in portions, and used as enzyme source for glycosyltransferase assays and Western blotting. Membrane bound recombinant $\beta 4$ -GlcNAcT was produced with a pMT2-based construct as described previously (3).

Western Blot Analysis—The Western blot analysis performed was aimed on detection of the protein A part of the fusion proteins. The proteins of 1 μ l of pooled medium collected after transfection, were separated by SDS-polyacrylamide gel electrophoresis on 10% gels using the Mini-PROTEAN II system (Bio-Rad). Western blotting was performed essentially as described previously (21). As first antibody, an arbitrary mouse IgG monoclonal antibody (ED3, Ref. 22) was used, which reacts with the protein A part of the hybrid proteins. The second antibody used was a goat anti-mouse peroxidase conjugate (Tago, Inc. Immunodiagnostic Reagents).

Glycosyltransferase Assays—Standard glycosyltransferase assays were performed in a 50- μ l reaction mixture containing either 25 nmol of UDP-[^3H]GalNAc (2 Ci/mol), UDP-[^3H]GlcNAc (1 Ci/mol), or UDP-[^3H]Gal (1 Ci/mol), 5 μ mol of sodium cacodylate buffer, pH 7.2, 1 μ mol of MnCl_2 , 0.2 μ mol of ATP, and 50 nmol of *p*-nitrophenyl-*N*-acetyl-1-thio- β -D-glucosaminide (GlcNAc-S-pNP). As enzyme source, 10 μ l of

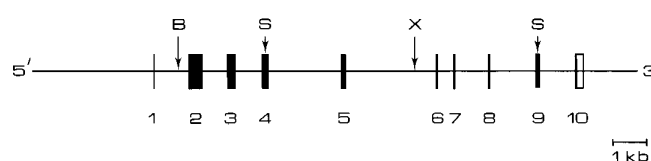


FIG. 1. Schematic representation of the *L. stagnalis* $\beta 4$ -GlcNAcT gene. Exons 1–10 are represented by vertical bars. B, S, and X indicate *Bam*HI, *Sal*I, and *Xho*I restriction sites, respectively.

COS cell medium, which had been concentrated 10 times with Centrprep-10 concentrators (Amicon), was used. The product was isolated using Sep-Pak C-18 cartridges (Waters) (23). For acceptor specificity studies, acceptor substrate concentrations were kept at 1 mM, in terms of terminal GlcNAc residues. Control assays lacking the acceptor substrate were carried out to correct for incorporation into endogenous acceptors. Enzyme activity was expressed as pmol/min $^{-1}$ /ml $^{-1}$ of the original medium. As enzyme source for the experiments studying the UDP-Gal specificity, the enzyme was isolated from the medium by binding to IgG-agarose (Sigma). 10 ml of medium was incubated with 10 μ l of IgG-agarose beads for 16 h at 4 °C, carefully shaking; the beads were subsequently collected by centrifugation and resuspended in 500 μ l of 0.1 M sodium cacodylate buffer, pH 7, containing 1 mg/ml bovine serum albumin; 10 μ l of this suspension was used in a standard assay.

Kinetic parameters (K_m and V) were estimated from Lineweaver-Burk plots by varying the sugar-donor concentrations from 0.05 to 0.5 mM for UDP-GlcNAc and from 0.25 to 5 mM for UDP-GalNAc while keeping the GlcNAc-S-pNP concentration at 1 mM, or by varying the acceptor substrate concentration from 0.025 to 1 mM while keeping the UDP-GlcNAc concentration at 0.5 mM. The inhibitory effect of UDP was studied at fixed concentrations of UDP-GlcNAc (0.5 mM) and GlcNAc-O-pNP (1 mM), whereas the UDP concentration was varied from 0.1 to 5 mM.

Product Characterization—Product, obtained with the enzyme protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8, GlcNAc-S-pNP, and UDP-[^3H]GalNAc, was analyzed by lectin affinity chromatography with *Wisteria floribunda* lectin (WFA), immobilized on azlactone/bisacrylamide polymeric beads (3 M EmphazeTM, Pierce) (24). The product was also analyzed by HPAEC-PAD (3). As a control, both WFA lectin chromatography and HPAEC-PAD analysis were performed with reference GalNAc $\beta 1 \rightarrow 4$ GlcNAc-S-pNP produced with *L. stagnalis* albumen gland $\beta 4$ -GalNAcT (9).

RESULTS

Isolation and Characterization of the *L. stagnalis* $\beta 4$ -GlcNAcT Gene—A genomic clone, denoted $\lambda 5$, was isolated previously from a λ EMBL3A library of *L. stagnalis*, and was shown to contain two short DNA sequences identical to $\beta 4$ -GlcNAcT cDNA sequences (3). A rough genomic map of $\lambda 5$ was constructed by PCR and Southern blot hybridization using specific $\beta 4$ -GlcNAcT cDNA fragments as probes (Fig. 1). $\lambda 5$ was found to encompass the complete coding sequence of the $\beta 4$ -GlcNAcT gene, spanning 12.5 kb of DNA, that was divided into 10 exons (Fig. 1, Table I). As probably part of the 5'-noncoding sequence is lacking from the cDNA (3), we cannot exclude the presence in the gene of one or more exons upstream of the denoted exon 1. Exon 10 was found to encompass the complete 3'-noncoding region. All exon sequences in the genomic clone were identical to those of the cDNA (3). Donor and acceptor splice junction sequences (Table II) are in agreement with consensus sequences reported (25).

Comparison of the *L. stagnalis* $\beta 4$ -GlcNAcT Gene with the Murine and Human $\beta 4$ -GalT Genes—The cDNA encoding $\beta 4$ -GlcNAcT shows sequence identity with the mammalian $\beta 4$ -GalT cDNAs identified (3, 5–7). A comparison of the protein-coding exons of the *L. stagnalis* $\beta 4$ -GlcNAcT gene with those of the murine and human $\beta 4$ -GalT genes (15, 16) is shown in Fig. 2. The $\beta 4$ -GalT gene was found to be divided into six exons, whereas the $\beta 4$ -GlcNAcT gene appeared to contain 10 exons. Exons 3, 4, 5, 6, and 9 of the $\beta 4$ -GlcNAcT gene were found to show similarity to exons 2–6 of the $\beta 4$ -GalT gene, corresponding to the catalytic domain of the enzyme (Fig. 2). This similarity was not only confined to a high degree of sequence

TABLE I
Exons of the *L. stagnalis* $\beta 4$ -GlcNAcT gene

Exon no.	Exon length bp	Corresponding cDNA residues	Amino acids encoded
1	>52	–3–52	1–17
2	399	53–451	18–150
3	233	452–684	151–228
4	191	685–875	229–292
5	123	876–998	293–333
6	102	999–1100	334–367
7	78	1101–1178	368–393
8	78	1179–1256	394–419
9	139	1257–1395	420–465
10	225	1396–1620	466–490

TABLE II
Introns of the *L. stagnalis* $\beta 4$ -GlcNAcT gene

Intron no.	Intron length kb	Splice junction sequence	
		Donor	Acceptor
1	1.0	AACAC:gtgagtagga	tttctcccag:GCCAT
2	0.7	CTTAG:gtataaaaaa	tttcgaacag:CTGGA
3	0.8	AACAG:gtgagaagga	gtcttgccag:ACTAC
4	2.1	TACAA:gtgagcatcc	tattttgcag:ACTTT
5	2.5	AACAG:gtaagacgtg	tcgattgcag:GGCCG
6	0.5	CACAG:gtgagaccca	atattaccag:CAAAG
7	0.9	AACAG:gtgagaccca	atattaccag:CAAA
8	1.3	GACAG:gtgagaccca	gtatccccag:ATTTA
9	1.0	CGAAT:gtgagtgttt	ctactttcag:AGCAT

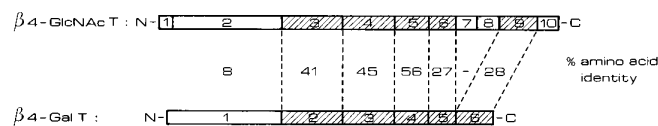


FIG. 2. Schematic representation of organization of the protein-coding exons of the genes encoding *L. stagnalis* $\beta 4$ -GlcNAcT and human $\beta 4$ -GalT (16). The shaded boxes indicate exons that show considerably amino acid identity between both genes. The open boxes indicate exons that show no significant sequence similarity between the genes, or are present only in the $\beta 4$ -GlcNAcT gene. Homologous splice sites are connected by dotted lines.

identity; intron/exon boundaries were also found at identical positions within the gene.

The coding sequence of the $\beta 4$ -GlcNAcT gene is larger than that of the $\beta 4$ -GalT gene. These additional sequences appeared to be mainly encoded by three exons (exons 7, 8, and 10), that were not present in the $\beta 4$ -GalT gene. Exons 7 and 8 were found to encode a partial repeat of the sequence of exon 6 (Table III). The sequence analysis of the intron/exon borders (Table II) shows that exons 6, 7, and 8 are symmetrical exons. Additionally, it was observed that the acceptor splice junction sites of introns 6 and 7 as well as the donor splice junction sites of intron 6, 7, and 8 and the adjacent exon sequences are identical. These data strongly suggest that exons 7 and 8 of the $\beta 4$ -GlcNAcT gene arose by duplications (26), originating from the downstream part of exon 6.

Construction and Expression of Deletion Derivatives of the $\beta 4$ -GlcNAcT cDNA—Deletion derivatives of the $\beta 4$ -GlcNAcT cDNA were constructed that lacked the sequences corresponding to exon 8 or to exons 7 and 8 (Fig. 3). Plasmids were constructed as gene fusions to part of protein A (20, 27). pMC 135 encodes a soluble, native $\beta 4$ -GlcNAcT (protA- $\beta 4$ -GlcNAcT). pMC142 and pMC166 are plasmids encoding deletion derivatives of pMC135, designated protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8 and protA- $\beta 4$ -GlcNAcT $\Delta 8$, respectively. These three plasmids were transiently expressed in COS cells. Essentially equal amounts of hybrid protein per volume were secreted into the medium

TABLE III
Exons of the $\beta 4$ -GlcNAcT gene encoding repeated amino acid sequences

Identical amino acids are indicated in boldface, and dash indicates a gap in the amino acid sequence.

Exon	Amino acid sequence encoded
6	AVHMKLPLLRKTL AHGLYDMVSH --- VEAGWNVNPHS
7	KGAHSLYDMLNKALGVQAGWNVHPNS
8	KWPLRLFLDSVNHAPAEAGAGWNVNPD

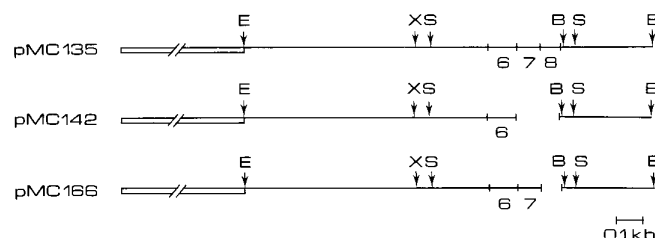


FIG. 3. Physical maps of the protein A- $\beta 4$ -GlcNAcT chimeric plasmids. The open box represents pPROTA vector, and the thin line inserted GlcNAcT cDNA. The positions of exons 6, 7, and 8 are indicated in the cDNA. E, X, S, and B indicate the position of the restriction sites for *EcoRI*, *XhoI*, *SalI*, and *BglII*, respectively, in the cDNA. pMC135 encodes protA- $\beta 4$ -GlcNAcT, pMC142 encodes protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8, and pMC166 encodes protA- $\beta 4$ -GlcNAcT $\Delta 8$.

(Fig. 4).

Enzyme Activity of the Chimeric Proteins—The enzyme activities of the mutant hybrid proteins and the native protA- $\beta 4$ -GlcNAcT were assayed using similar amounts of enzyme (Fig. 4). As the native membrane-bound $\beta 4$ -GlcNAcT shows a low GalNAcT activity (about 7% of the $\beta 4$ -GlcNAcT activity), the activity of the enzymes was measured with both UDP-GlcNAc and UDP-GalNAc. ProtA- $\beta 4$ -GlcNAcT $\Delta 7$ –8 showed reproducibly an almost 2 times higher GlcNAcT activity and a 4 times higher GalNAcT activity than the parental protA- $\beta 4$ -GlcNAcT (Table IV). In contrast to the other chimeric proteins, protA- $\beta 4$ -GlcNAcT $\Delta 8$ appeared to be enzymatically inactive. To determine if the enzymes would show GalT activity, the hybrid proteins were purified using IgG-agarose beads. In this way, the recombinant enzymes were completely disposed of COS cell-derived $\beta 4$ -GalT. The bead-associated chimeras did not show detectable GalT activity, whereas they showed GlcNAcT and GalNAcT activities similar to those for the concentrated media (results not shown).

Acceptor specificity studies, using UDP-GlcNAc as sugar donor, showed no significant differences in acceptor preference between the two active hybrid enzymes at an acceptor concentration of 1 mM (Table V; Ref. 4). The acceptor specificity of the mutant chimeric protein using UDP-GalNAc as a sugar-donor appeared very similar to the preference of the enzyme when using UDP-GlcNAc. By contrast, the prostate gland $\beta 4$ -GalNAcT is less specific for the linkage type of the terminal β -linked GlcNAc, and its acceptor substrate requirement resembles that of the albumen gland $\beta 4$ -GalNAcT (9).

Kinetic Properties of the Chimeric $\beta 4$ -GlcNAcTs—To explain the differences in donor specificity in more detail, the K_m and V values for UDP-GlcNAc and UDP-GalNAc were determined for both protA- $\beta 4$ -GlcNAcT and protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8. It appears that the increase in GlcNAcT activity of protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8 is due to a 3-fold reduction in K_m for UDP-GlcNAc (Table VI), whereas the increased GalNAcT activity of the mutant enzyme is mainly due to an enhanced maximum velocity. These effects result in an elevated kinetic efficiency for both types of transfer. The data suggest an involvement of the region encoded by exons 7 and 8 in UDP-sugar donor binding. However, it was found that both active chimera were

inhibited to almost the same extent by UDP (50% at 4–5 mM UDP, results not shown).

The K_m and V values of two acceptor substrates were estimated (Table VII). While the V values of both enzymes appeared to be of the same order for both compounds, the K_m values were decreased in the mutant enzyme. Deletion of the two exons clearly elevates the kinetic efficiency with both acceptor substrates.

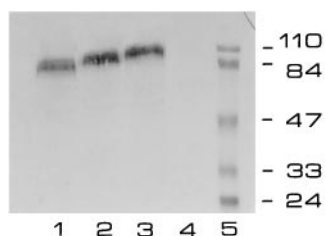


FIG. 4. **Western blot analysis of protA-fusion proteins.** COS-7M6 cells were transfected with the pPROTA vector and the pPROTA-fusion plasmids, carrying the $\beta 4$ -GlcNAcT cDNA, or deletion derivatives. Total cellular proteins, secreted into the medium, were separated by SDS-polyacrylamide gel electrophoresis (10%). The proteins were transferred to nitrocellulose. Antiserum reactions were directed against the protein A part of the proteins. Lane 1, protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8; lane 2, protA- $\beta 4$ -GlcNAcT $\Delta 8$; lane 3, protA- $\beta 4$ -GlcNAcT; lane 4, pPROTA; lane 5, molecular mass standards. Molecular masses on the right are indicated in kDa.

TABLE IV
UDP-sugar donor promiscuity of protein A- $\beta 4$ -GlcNAcT
chimeras and deletion mutants derived thereof

Glycosyltransferase activities of the concentrated medium of transfected COS cells were assayed as described under "Experimental Procedures" using UDP-GlcNAc and UDP-GalNAc as donors, respectively.

Enzyme	Glycosyltransferase activity with	
	UDP-GlcNAc	UDP-GalNAc
	<i>pmol · min⁻¹ · ml⁻¹</i>	
protA- $\beta 4$ -GlcNAcT	210	15
protA- $\beta 4$ -GlcNAcT $\Delta 8$	<1	<1
protA- $\beta 4$ -GlcNAcT $\Delta 7,8$	361	67

TABLE V
Acceptor specificity of recombinant protA- $\beta 4$ -GlcNAcT and protA- $\beta 4$ -GlcNAcT $\Delta 7,8$

For the GlcNAcT assays 100% activity is 128 and 338 $\text{pmol} \cdot \text{min}^{-1} \cdot \text{ml}^{-1}$ medium⁻¹, respectively. For the GalNAcT assays 100% activity is 110 $\text{pmol} \cdot \text{min}^{-1} \cdot \text{ml}^{-1}$ medium⁻¹ for the mutant enzyme, and 5 $\text{nmol} \cdot \text{min}^{-1} \cdot \text{ml}^{-1}$ for the prostate-derived enzyme.

Acceptor substrate	Activity with				
	UDP-GlcNAc			UDP-GalNAc	
	protA- $\beta 4$ -GlcNAcT	protA- $\beta 4$ -GlcNAcT $\Delta 7,8$	rec- $\beta 4$ -GlcNAcT ^a	protA- $\beta 4$ -GlcNAcT $\Delta 7,8$	Prostate gland membranes
			%		
1 GlcNAc β -S-pNP	100	100	100	100	100
2 GlcNAc β -O-pNP	55	72	71	68	28
3 GlcNAc $\beta 1 \rightarrow 4$ GlcNAc β -O-pNP	5	7	1	1	11
4 GlcNAc $\beta 1 \rightarrow 2$ Man $\alpha 1 \rightarrow 6$ GlcNAc $\beta 1 \rightarrow 2$ Man $\alpha 1 \rightarrow 3$ Man β -O-R ^b	<1	<1	<1	1	48
5 GlcNAc $\beta 1 \rightarrow 3$ GalNAc α -O-pNP	3	4	5	2	47
6 GlcNAc $\beta 1 \rightarrow 6$ GalNAc α -O-pNP Gal $\beta 1 \rightarrow 3$	96	86	104	136	67

^a Recombinant soluble $\beta 4$ -GlcNAcT from baculovirus infected insect cells (4).

^b R = $-(\text{CH}_2)_8\text{-COOCH}_3$.

TABLE VI
Kinetic parameters of protein A- $\beta 4$ -GlcNAcT chimeras and a deletion mutant derived thereof for two UDP-sugar donors

The recombinant enzymes were produced as protein A fusion proteins in COS cells.

Enzyme	UDP-GlcNAc			UDP-GalNAc		
	K_m	V	Kinetic efficiency	K_m	V	Kinetic efficiency
	<i>mM</i>	<i>pmol · min⁻¹ · ml⁻¹</i>	<i>V · K_m⁻¹</i>	<i>mM</i>	<i>pmol · min⁻¹ · ml⁻¹</i>	<i>V · K_m⁻¹</i>
protA- $\beta 4$ -GlcNAcT	0.23 ± 0.02	300 ± 11	1304	2.0 ± 0.6	92 ± 13	46
protA- $\beta 4$ -GlcNAcT $\Delta 7,8$	0.08 ± 0.01	380 ± 12	4750	1.6 ± 0.2	358 ± 19	224

Characterization of the N-Acetylgalactosaminylated Product Obtained with $\beta 4$ -GlcNAcT $\Delta 7$ –8—The N-acetylgalactosaminylated product formed by incubating protA- $\beta 4$ -GlcNAcT $\Delta 7$ –8, UDP-GalNAc, and GlcNAc-S-pNP was anticipated to be GalNAc $\beta 1 \rightarrow 4$ GlcNAc-S-pNP. To confirm this, the product was compared with authentic GalNAc $\beta 1 \rightarrow 4$ GlcNAc-S-pNP obtained by the action of *L. stagnalis* albumen gland $\beta 4$ -GalNAcT (9). Both product and reference were subjected to lectin affinity chromatography with immobilized WFA, known to bind with high affinity to glycans containing terminal GalNAc in a $\beta 1 \rightarrow 4$ -linkage (24, 28, 29). More than 90% of both compounds bound to the immobilized WFA, and could be eluted with 10 mM GalNAc. In addition, reference and product eluted on HPAEC with the same retention time (data not shown). These results indicate that $\beta 4$ -GlcNAcT $\Delta 7$ –8 shows a $\beta 4$ -GalNAcT activity.

DISCUSSION

Based on the sequence identity of the mammalian $\beta 4$ -GalTs and *L. stagnalis* $\beta 4$ -GlcNAcT, these enzymes have been proposed to be members of one gene family (3). The observed conservation of the exon-intron organization of the genes encoding these enzymes strengthens this view. All exons of the $\beta 4$ -GalT gene show sequence identity to parts of the $\beta 4$ -GlcNAcT gene, which suggests that both genes originated from a common ancestor, and diverged to genes encoding enzymes with a different sugar donor specificity.

The most remarkable difference between the genes is found in an insertion of two exons (exons 7 and 8) in the $\beta 4$ -GlcNAcT gene. These exons were found to encode partial repeats of exon 6. The sequences of these exons and those of the bordering introns strongly suggest that exons 7 and 8 were generated by internal exon duplications, and originate from exon 6. This is the first glycosyltransferase described in which exon duplication seems to have taken place during evolution. Generally, exon duplication is thought to play an important role in the evolution of genes, and many complex genes have been described that have evolved by internal exon duplication and subsequent modification of the primordial genes (30, 31).

TABLE VII

Kinetic parameters of protein A- $\beta 4$ -GlcNAcT chimeras and a deletion mutant derived thereof for two acceptor substrates

The recombinant enzymes were produced as protein A fusion proteins in COS cells.

Enzyme	GlcNAc β -O-pNP			GlcNAc $\beta 1 \rightarrow 6$ Gal $\beta 1 \rightarrow 3$ GalNAc α -O-pNP		
	K_m	V	Kinetic efficiency	K_m	V	Kinetic efficiency
	mM	nmol \cdot min $^{-1}$ \cdot ml $^{-1}$	V \cdot K $_m^{-1}$	mM	nmol \cdot min $^{-1}$ \cdot ml $^{-1}$	V \cdot K $_m^{-1}$
protA- $\beta 4$ -GlcNAcT	0.32 \pm 0.03	1.53 \pm 0.07	5	0.09 \pm 0.01	1.84 \pm 0.05	20
protA- $\beta 4$ -GlcNAcT $\Delta 7-8$	0.13 \pm 0.01	1.66 \pm 0.03	13	0.02 \pm 0.005	1.16 \pm 0.05	58

To study the effect of the exon duplications on enzyme catalysis, we deleted the sequence (52 amino acids) corresponding to the two additional exons from the cDNA sequence. The catalytic domain of the mutant protein thus obtained, might resemble that of a putative ancestor of the *L. stagnalis* $\beta 4$ -GlcNAcT. The deletion enhanced the kinetic efficiency of the enzyme for both the transfer of GlcNAc and GalNAc, as well as for its acceptor substrates. The acceptor substrate specificity, i.e. the preference for a terminal $\beta 6$ -linked GlcNAc, was not affected. So the mutant enzyme has an enhanced catalytic potential, but also an increased sugar-donor promiscuity. Surprisingly, deletion of only the sequence encoded by exon 8 resulted in a completely inactive enzyme. An explanation could be that the enzyme is not properly folded with only one additional repeat, whereas the presence of two additional copies of the repeated sequence allows a correct folding.

From the results obtained here, it is difficult to deduce the selective advantage that was obtained by the exon duplications for the biological function of the $\beta 4$ -GlcNAcT. As can be inferred from the sequence, changes have occurred in the repeats after the exon duplications, and most likely also in other regions of the protein. It is possible, however, that reduction of the capacity of the primordial $\beta 4$ -GlcNAcT to transfer GalNAc might have been advantageous for the snail. The increase in the specificity of the enzyme for UDP-GlcNAc would then have been of more importance for the snail than the loss of catalytic potency that coincided with the introduction of the additional exons.

In $\beta 4$ -GalT the corresponding region (encoded by exon 5) has been proposed to be involved in UDP-Gal binding (32–34). In the same region a tetrapeptide (DKKN) has been found that is conserved between $\beta 4$ -GalT and $\alpha 3$ -GalT, which is in support of this proposition (35). In the human $\beta 4$ -GalT a second region (in exon 4) has been proposed to be involved in UDP-Gal binding (36, 37). Interestingly, this latter region shows a high degree of similarity between the $\beta 4$ -GalTs and the $\beta 4$ -GlcNAcT, whereas the more downstream region of $\beta 4$ -GalT (encoded by exon 5) shows much less sequence identity with the corresponding region in the $\beta 4$ -GlcNAcT (encoded by exons 6, 7, and 8). As the mammalian $\beta 4$ -GalTs and the *L. stagnalis* $\beta 4$ -GlcNAcT use UDP-Gal and UDP-GlcNAc, respectively, it is tempting to assume that the domain that is most conserved between these enzymes, is involved in the interaction with the UDP part of the nucleotide-sugar, while the more downstream region is responsible for the specificity for donor Gal and GlcNAc, respectively. The observation that both native $\beta 4$ -GlcNAcT and the deletion mutant were inhibited to a similar extent by UDP, suggesting that the UDP-binding domain is not affected by the deletion, is in support of this supposition. The lower K_m for UDP-GlcNAc that was found for the mutant enzyme could be explained by a higher affinity of this enzyme for the GlcNAc part of the UDP-sugar.

The change in kinetic properties observed with the mutant enzyme suggests that the region around the insertion in $\beta 4$ -GlcNAcT is not only involved in interaction with the sugar-donor, but also in binding of the acceptor substrate. Addition-

ally, in $\beta 4$ -GalT a region (in exon 4) has been identified, that is involved in interaction of both donor and acceptor substrates (36, 37). So in both enzymes the donor and acceptor substrate binding domains seem to be structurally linked. This is conceivable, as both substrates have to be close together for the transfer reaction. In $\beta 4$ -GalT another binding domain for GlcNAc, however, has been localized in exon 2/3 (34). This region shows a high sequence identity with the corresponding region in $\beta 4$ -GlcNAcT, which supports the suggestion that this sequence is involved in acceptor binding.

We have shown here that the *L. stagnalis* $\beta 4$ -GlcNAcT can be transformed to an enzyme that is capable of catalyzing the transfer of GlcNAc and GalNAc with a similar maximum velocity. This suggests that the $\beta 4$ -GalNAcT, which has been observed in *L. stagnalis* tissues (Ref. 9 and this study), might be encoded by a structurally related enzyme belonging to the $\beta 4$ -GalT gene family. This is further supported by several studies that have shown a similarity between invertebrate $\beta 4$ -GalNAcTs and the mammalian $\beta 4$ -GalTs in acceptor specificity (9–12), and sometimes in responsiveness to α -lactalbumin (13). Furthermore, $\beta 4$ -GalT shows sugar donor promiscuity at high concentrations of UDP-GalNAc (38), or in the presence of α -lactalbumin (24), resulting in the transfer of GalNAc. Interestingly, $\beta 4$ -GlcNAcT is not sensitive to α -lactalbumin (4), and this enzyme can not utilize UDP-Gal, but shows a relatively high $\beta 4$ -GalNAcT activity (7%). By mutagenesis, we have constructed an enzyme with an even higher GalNAcT activity. A similar experiment has been documented for the blood group A ($\alpha 3$ -GalNAcT) and B ($\alpha 3$ -GalT) enzymes (39). These enzymes also show some nucleotide sugar donor promiscuity (40), but by construction of hybrids an enzyme with both activities was obtained (41). Our observations suggest that, in addition to the facilitation of the transfer of the desired sugar, the prevention of sugar donor promiscuity might have been a driving force in the evolution of enzymes of the $\beta 4$ -GalT gene family.

Acknowledgments—We thank J. Aino Andriessen, Casper Kee, and Carolien Koeleman for technical assistance in part of the experiments; Dr. Timo K. van den Berg for the gift of monoclonal antibody ED3; Dr. Bruce A. Macher for the gift of pPROTA vector; Dr. Ole Hindsgaul for the gift of oligosaccharide compound 4; and Dr. Richard D. Cummings for advice on lectin chromatography and the gift of WFA-Emphase™.

REFERENCES

- Schachter, H. (1991) *Curr. Opin. Struct. Biol.* **1**, 755–765
- Van den Eijnden, D. H., and Joziassse, D. H. (1993) *Curr. Opin. Struct. Biol.* **3**, 711–721
- Bakker, H., Agterberg, M., Van Tetering, A., Koeleman, C. A. M., Van den Eijnden, D. H., and Van Die, I. (1994) *J. Biol. Chem.* **269**, 30326–30333
- Bakker, H., Schoenmakers, P. S., Koeleman, C. A. M., Joziassse, D. H., Van Die, I., and Van den Eijnden, D. H. (1997) *Glycobiology* **7**, 539–548
- Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R., and Shaper, J. H. (1988) *J. Biol. Chem.* **263**, 10420–10428
- Shaper, N. L., Shaper, J. H., Meuth, J. L., Fox, J. L., Chang, H., Kirsch, I. R., and Hollis, G. F. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1573–1577
- Masri, K. A., Appert, H. E., and Fukuda, M. N. (1988) *Biochem. Biophys. Res. Commun.* **157**, 657–663
- Van den Eijnden, D. H., Neeleman, A. P., Van der Knaap, W. P. W., Bakker, H., Agterberg, M., and Van Die, I. (1995) *Adv. Exp. Med. Biol.* **376**, 47–51
- Mulder, H., Spronk, B. A., Schachter, H., Neeleman, A. P., Van den Eijnden, D. H., De Jong-Brink, M., Kamerling, J. P., and Vliegthart, J. F. G. (1995) *Eur. J. Biochem.* **227**, 175–185
- Neeleman, A. P., Van der Knaap, W. P. W., and Van den Eijnden, D. H. (1994)

- Glycobiology* **4**, 641–651
11. Van Die, I., Van Tetering, A., Bakker, H., Van den Eijnden, D. H., and Joziassse, D. H. (1996) *Glycobiology* **6**, 157–164
 12. Srivatsan, J., Smith, D. F., and Cummings, R. D. (1994) *J. Parasitol.* **80**, 884–890
 13. Neeleman, A. P., and Van den Eijnden, D. H. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10111–10116
 14. Van den Eijnden, D. H., Neeleman, A. P., Van der Knaap, W., Bakker, H., Agterberg, M., and Van Die, I. (1995) *Biochem. Soc. Trans.* **23**, 175–179
 15. Hollis, G. F., Douglas, J. G., Shaper, N. L., Shaper, J. H., Stafford-Hollis, J. M., Evans, R. J., and Kirsch, I. R. (1989) *Biochem. Biophys. Res. Commun.* **162**, 1069–1075
 16. Mengle-Gaw, L., McCoy-Haman, M. F., and Tiemeier, D. C. (1991) *Biochem. Biophys. Res. Commun.* **176**, 1269–1276
 17. Bogerd, J., Vankesteren, R. E., Vanheerikhuizen, H., Geraerts, W. P. M., Veenstra, J., Smit, A. B., and Joosse, J. (1993) *Cell Mol. Neurobiol.* **13**, 123–136
 18. Del Sal, G., Manfioletti, G., and Schneider, C. (1988) *Nucleic Acids Res.* **16**, 9878
 19. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
 20. Sanchez-Lopez, R., Nicholson, R., Gesnel, M.-C., Matrisian, L. M., and Breathnach, R. (1988) *J. Biol. Chem.* **263**, 11892–11899
 21. Agterberg, M., Van Die, I., Yang, H., Andriessen, J. A., Van Tetering, A., Van den Eijnden, D. H., and Ploegh, H. L. (1993) *Eur. J. Biochem.* **217**, 241–246
 22. Dijkstra, C. D., Döpp, E. A., Joling, P., and Kraal, G. (1985) *Immunology* **54**, 589
 23. Palcic, M. M., Heerze, L. D., Pierce, M., and Hindsgaul, O. (1988) *Glycoconjugate J.* **5**, 49–63
 24. Do, K.-Y., Do, S.-I., and Cummings, R. D. (1995) *J. Biol. Chem.* **270**, 18447–18451
 25. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383
 26. Patthy, L. (1994) *Curr. Opin. Struct. Biol.* **4**, 383–392
 27. Uhlen, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L., and Lindberg, M. (1984) *J. Biol. Chem.* **259**, 1695–1702
 28. Torres, B. V., McCrumb, D. K., and Smith, D. F. (1988) *Arch. Biochem. Biophys.* **262**, 1–11
 29. Nyame, K., Smith, D. F., Damian, R. T., and Cummings, R. D. (1989) *J. Biol. Chem.* **264**, 3235–3243
 30. Li, W.-H. (1983) in *Evolution of Genes and Proteins* (Nei, M., and Koehn, R. K., eds) pp. 14–37, Sinauer Associates, Sunderland, MA
 31. Li, W.-H., and Graur, D. (1991) *Fundamentals of Molecular Evolution*, Sinauer Associates, Sunderland, MA
 32. Yadav, S., and Brew, K. (1990) *J. Biol. Chem.* **265**, 14163–14169
 33. Wang, Y., Wong, S. S., Fukuda, M. N., Zu, H. Y., Liu, Z. D., Tang, Q. S., and Appert, H. E. (1994) *Biochem. Biophys. Res. Commun.* **204**, 701–709
 34. Boeggeman, E. E., Balaji, P. V., and Qasba, P. K. (1995) *Glycoconjugate J.* **12**, 865–878
 35. Joziassse, D. H., Shaper, J. H., Van den Eijnden, D. H., Van Tunen, A. J., and Shaper, N. L. (1989) *J. Biol. Chem.* **264**, 14290–14297
 36. Aoki, D., Appert, H. E., Johnson, D., Wong, S. S., and Fukuda, M. N. (1990) *EMBO J.* **9**, 3171–3178
 37. Zu, H., Fukuda, M. N., Wong, S. S., Wang, Y., Liu, Z., Tang, Q., and Appert, H. E. (1995) *Biochem. Biophys. Res. Commun.* **206**, 362–369
 38. Palcic, M. M., and Hindsgaul, O. (1991) *Glycobiology* **1**, 205–209
 39. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345**, 229–233
 40. Watkins, W. M., Greenwell, P., Yates, A. D., and Johnson, P. H. (1988) *Biochimie* **70**, 1597–1611
 41. Yamamoto, F., and Hakomori, S. (1990) *J. Biol. Chem.* **265**, 19257–19262